

PARAMETER ESTIMATION FOR GRAVITATIONAL-WAVE BURSTS WITH THE BAYESWAVE PIPELINE

BENCE BÉCSY^{1,2}, PETER RAFFAI^{1,2}, NEIL J. CORNISH³, REED ESSICK⁴, JONAH KANNER⁵, ERIK KATSAVOUNIDIS⁴, TYSON B. LITTENBERG⁶, MARGARET MILLHOUSE³, AND SALVATORE VITALE⁴

¹Institute of Physics, Eötvös University, 1117 Budapest, Hungary; becsybence@caesar.elte.hu

²MTA-ELTE EIRSA “Lendület” Astrophysics Research Group, 1117 Budapest, Hungary

³Department of Physics, Montana State University, Bozeman, MT 59717, USA

⁴Massachusetts Institute of Technology, 185 Albany St, 02139 Cambridge USA

⁵LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA

⁶NASA Marshall Space Flight Center, Huntsville AL 35812, USA

ABSTRACT

We provide a comprehensive multi-aspect study on the performance of a pipeline used by the LIGO-Virgo Collaboration for estimating parameters of gravitational-wave bursts. We add simulated signals with four different morphologies (sine-Gaussians, Gaussians, white-noise bursts, and binary black hole signals) to simulated noise samples representing noise of the two Advanced LIGO detectors during their first observing run. We recover them with the BayesWave (BW) pipeline to study its accuracy in sky localization, waveform reconstruction, and estimation of model-independent waveform parameters. BW localizes sources with a level of accuracy comparable for all four morphologies, with the median separation of actual and estimated sky locations ranging from 25.1° to 30.3°. This is a reasonable accuracy in the two-detector case, and is comparable to accuracies of other localization methods studied previously. As BW reconstructs generic transient signals with sine-Gaussian wavelets, it is unsurprising that BW performs the best in reconstructing sine-Gaussian and Gaussian waveforms. BW’s accuracy in waveform reconstruction increases steeply with network signal-to-noise ratio (SNR_{net}), reaching a 85% and 95% match between the reconstructed and actual waveform below $\text{SNR}_{\text{net}} \approx 20$ and $\text{SNR}_{\text{net}} \approx 50$, respectively, for all morphologies. BW’s accuracy in estimating waveform moments is dominated by its accuracy in waveform reconstruction, and generally higher order moments can be estimated with lower accuracy. The figures of merit we introduce can be used in future characterizations of parameter estimation pipelines, and allow fair comparisons between different methods.

Keywords: gravitational waves – methods: data analysis

1. INTRODUCTION

The network of Advanced LIGO (aLIGO) gravitational-wave (GW) detectors ([Aasi et al. 2015](#)), consisting of aLIGO-Hanford (H1) and aLIGO-Livingston (L1), finished its first observing run (O1) in January 2016. During O1, this network achieved the first direct detections of GWs by detecting GW150914 ([Abbott et al. 2016d](#)) and GW151226 ([Abbott et al. 2016b](#)), two signals from coalescences of binary black holes. Besides binary black holes, other astrophysical sources of GW transients (e.g. core-collapse supernovae, magnetar flares, and cosmic string cusps) are also targeted by aLIGO ([Abbott et al. 2016g](#)). Searches for generic GW transients aim to detect weakly-modeled GW signals (“bursts”) from such systems, as well as from binary black holes, and also from as-yet-unknown sources (see e.g. [Abbott et al. 2016e](#)).

Detections of GW signals will be used to test and constrain models of astrophysical sources (see e.g. [Abbott et al. 2016a](#)). This usually requires reconstructing the signal waveform from the GW detector output and estimating parameters of the waveform (see e.g. [Abbott et al. 2016f](#)). For sources where an accurate waveform model exists, such as binary black holes in circular orbits, this is done by matching the detector output with template waveforms (see e.g. [Abbott et al. 2016f](#)). In this case, the estimated parameters are astrophysical, e.g., chirp mass and spins. Parameter estimation (PE) for burst signals, where no model templates exist, need a different approach. In such cases, basis functions are used to reconstruct the waveform and to estimate model-independent parameters of it, such as central time and frequency, signal duration and bandwidth. Besides these intrinsic parameters of the wave-

form, estimates can also be given on the extrinsic parameters of the source (e.g. sky location).

BayesWave (BW) is a pipeline for detecting and characterizing GW bursts, that works within the framework of Bayesian statistics and uses sine-Gaussian wavelets as basis functions to reconstruct the signal (Cornish & Littenberg 2015). In O1, BW was used as a follow-up PE tool on triggers provided by the coherent Waveburst (cWB) search pipeline (Klimenko et al. 2008, 2016), which identifies coincident excess power in strain data of multiple GW detectors. Note however, that cWB can also reconstruct the sky location of a GW source and the waveform of the GW signal, independently from BW (Klimenko et al. 2011). This provides an opportunity to compare the performances of BW and cWB in PE using the same set of triggers (for the results of this comparison, see Section 3.1). BW is effective in distinguishing GW signals from non-Gaussian noise artifacts (“glitches”), which enables the combination of the cWB and BW pipelines to achieve high-confidence detections across a range of waveform morphologies (Littenberg et al. 2016; Kanner et al. 2016). The estimates of mass parameters and sky location obtained by BW for GW150914 have shown to be consistent with template-based PE pipelines (Abbott et al. 2016e).

In this paper we characterize BW’s performance in PE by injecting a large set of simulated signals into simulated aLIGO noise, and recovering them and their parameters with BW. The main purpose of this study is to determine the accuracy of these reconstructions that can be achieved with BW. By knowing the accuracy, future studies can identify the broadest range of astrophysical models that can be tested with BW, while further improvements of BW can be guided by these results. Among the estimated parameters, we give special attention to sky location of the GW source, because of its key role in electromagnetic (EM) follow-up observations of GW events (see e.g. Abbott et al. 2016c; Singer et al. 2014; Berry et al. 2015; Vitale et al. 2016). Sky localization of GW burst sources can be carried out with the cWB and LALInferenceBurst (LIB) pipelines (Lynch et al. 2015; Veitch et al. 2015) too. An extensive analysis on the sky localization performance of cWB and LIB was published in Essick et al. (2015). Here we present a similar analysis for BW in order to characterize its performance and to allow comparisons with other burst pipelines studied in Essick et al. (2015). Note however, that as we use a reduced set of triggers compared to Essick et al. (2015) (for an explanation, see Section A in the Appendix), our results in Figures 1-4 should not be compared directly with results in Figures 3-6 of Essick et al. (2015). Instead, to allow direct comparisons between BW, cWB, and LIB, we repeat

our analysis with cWB and with LIB on the same reduced set of triggers, and present the results in Figure Sets 1-4 (available in the online journal). Also note that new cWB sky localization results for binary black holes presented recently (see Vitale et al. 2016) show that cWB’s performance has improved significantly for a three-detector network, while it has not changed significantly for the two-detector case we present here.

We focus on three aspects of BW’s performance: (i) sky localization, (ii) waveform reconstruction, and (iii) estimation of model-independent waveform parameters. In Section 2, we describe the methods used for creating simulated signals and noise samples, and used by BW for carrying out PE. In Section 3, we present results of our analyses regarding all (i)-(iii) aspects. We summarize our findings and highlight some implications in Section 4.

2. METHODS

We used software injections to test the PE performance of BW, i.e. we created mock samples of aLIGO noise and added simulated GW signals with four different morphologies to these samples. We then used these samples at trigger times provided by cWB as inputs for BW to test what it recovers from the signals embedded in the mock detector noise. In this section we discuss the characteristics of noise samples and of simulated signals we used (Section 2.1), as well as methods BW uses for PE (Section 2.2).

2.1. Noise and injections

In our analysis we considered a two-detector network consisting of H1 and L1. We used the same stationary and Gaussian mock noise samples as in the study presented in Essick et al. (2015). These samples were generated using the expected 2015 sensitivity curve of aLIGO (see the top panel of Fig. 1 in Essick et al. 2015), thus they have slightly different characteristics than the actual noise collected during the O1 run. Projections show that the two LIGO detectors will operate in the first two months of the second observing run (O2) with similar sensitivity curves they operated with during O1. Thus, we expect that our results are representative for this first period of O2 as well.

We used the same set of signals in our software injections as in the Essick et al. (2015) study. This set consists of signals with four different morphologies: sine-Gaussians (SG); Gaussians (G); white-noise bursts (WNB); and binary black hole (BBH) mergers. This wide range of signal morphologies allows us to test the PE performance of BW with minimal assumptions on the GW signal. Parameters of injected signals were randomly drawn from distributions shown in Table 4 of Essick et al. (2015). The amplitude distribution of in-

jected signals was chosen such as to represent a uniform distribution of GW sources in volume (see Appendix C in [Essick et al. 2015](#)). Signal injections were distributed uniformly over the sky and were regularly spaced in time.

Since in present searches BW only runs on triggers provided by cWB ([Abbott et al. 2016e](#)), the number of signals we analyzed with BW for each morphology was determined by the cWB search on the whole data sample. In our analysis, cWB found 1112 *SG*, 256 *G*, 769 *WNB*, and 2488 *BBH* from all injected signals. Note that similarly to the study presented in [Essick et al. \(2015\)](#), we only used 500 of the *BBH* triggers in order to reduce computational costs.

SG waveforms are often used to model generic transients (e.g. [Abadie et al. 2012](#)), because they are the most localized signals in time-frequency space where generic burst searches (including cWB) operate (see [Chatterji 2005](#)). We define *SG* waveforms with the following two equations:

$$h_+(t) = \cos(\alpha)h_{\text{rss}}\sqrt{\frac{4f_0\sqrt{\pi}}{Q(1 + \cos(2\phi_0)e^{-Q^2})}} \times \cos(2\pi f_0(t - t_0) + \phi_0)e^{-(t-t_0)^2/\tau^2} \quad (1a)$$

$$h_\times(t) = \sin(\alpha)h_{\text{rss}}\sqrt{\frac{4f_0\sqrt{\pi}}{Q(1 - \cos(2\phi_0)e^{-Q^2})}} \times \sin(2\pi f_0(t - t_0) + \phi_0)e^{-(t-t_0)^2/\tau^2}, \quad (1b)$$

where $\alpha \in [0, \pi/2]$ is a parameter that sets the relative weights between polarizations h_+ and h_\times , $h_{\text{rss}}^2 = \int (h_+^2 + h_\times^2) dt$ is the root-sum-squared strain amplitude chosen as a free parameter in the amplitude randomization process, f_0 is the central frequency, t_0 is the central time, ϕ_0 is the phase at time $t = t_0$, τ is the width of the signal in the time domain, and $Q = \sqrt{2}\pi\tau f_0$ is the quality factor encoding the characteristic number of cycles within duration of the signal.

G signals are the special cases of *SG* signals when $f_0 \rightarrow 0$, and are defined as:

$$h_+(t) = \cos(\alpha)\frac{h_{\text{rss}}}{\sqrt{\tau}}\left(\frac{2}{\pi}\right)^{1/4}e^{-(t-t_0)^2/\tau^2} \quad (2a)$$

$$h_\times(t) = \sin(\alpha)\frac{h_{\text{rss}}}{\sqrt{\tau}}\left(\frac{2}{\pi}\right)^{1/4}e^{-(t-t_0)^2/\tau^2}. \quad (2b)$$

Despite their similarity to *SG*s, these signals pose different challenges, because they have their highest amplitude at $f = 0$ Hz in frequency domain, and thus they have most of their power at low frequencies where aLIGO is less sensitive.

WNB waveforms are intended to model a time-localized excess power uniformly distributed in a given frequency band, and satisfy:

$$h_{+,\times}(t) \propto e^{\frac{-(t-t_0)^2}{\tau^2}} \int_{-\infty}^{\infty} e^{-i2\pi ft} w(f) df, \quad (3)$$

where $w(f)$ values are randomly drawn from a Gaussian white noise within and chosen to be $w(f) = 0$ outside the band $f \in [f_{\min}, f_{\max}]$. We generated the right side of equation (3) independently for the $+$ and \times polarizations, and normalized them to get h_+ and h_\times with the desired h_{rss} . Unlike signals with the other three morphologies, *WNB* signals are not elliptically polarized, because the procedure used to produce them generates h_+ and h_\times independently.

The only astrophysical signals we used were *BBH*s with spins aligned or anti-aligned with the orbital angular momentum. We only considered binaries with relatively high detector-frame total masses ($M_{\text{tot}} \in [30, 50] M_\odot$), because their signals are more compact in time-frequency space, which makes them good targets for generic burst searches. Three different methods have been used for calculating the waveform in the three different phases of binary evolution: 3.5PN post-Newtonian expansion, numerical relativity, and analytic quasi-normal modes to calculate the inspiral, merger, and ringdown waveforms, respectively (see [Ajith et al. 2011](#) and [Hannam et al. 2010](#) for details).

2.2. The BayesWave pipeline

BW uses a trans-dimensional Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm ([Green 1995](#)) to explore the following three competing models of the data, and test them with the input data samples from each aLIGO detector: i) Gaussian noise only; ii) Gaussian noise with glitches; iii) Gaussian noise with a GW signal. This approach makes BW effective in distinguishing GW signals from glitches ([Littenberg et al. 2016](#)), but it also makes BW computationally expensive, and thus in O1 BW was used to follow-up candidate events from cWB.

BW assumes that all signals are elliptically polarized, i.e. $h_\times = \epsilon h_+ e^{i\pi/2}$, where $\epsilon \in [0, 1]$ is the ellipticity parameter, which is 0 for linearly polarized signals and 1 for circularly polarized ones. This is a valid assumption for many expected astrophysical signals, but not for our injections with *WNB* morphology (see Section 2.1). However, for a LIGO-only network, it is often the case that only a combination of the two polarizations, rather than the separate $+$ and \times components, will be detectable, making the elliptical constraint a fair approximation for many cases.

We used the BW version which had been used for the offline analysis of O1 data to attain a characterization

of BW’s performance during O1, and to support a fair comparison with the versions of other PE pipelines characterized in [Essick et al. \(2015\)](#). PE pipelines used by the LIGO-Virgo Collaboration (including BW) have undergone improvements since the beginning of O1 (some of which were motivated by this study).

3. RESULTS

In this section we show how BW performed in different aspects of PE. These aspects are sky localization (see Section 3.1), waveform reconstruction (see Section 3.2), and point estimates of waveform moments (see Section 3.3).

We used the same version of BW which was used during O1. It identified 779 SG, 355 WNB, and 1 BBH injections as glitches or Gaussian noise only, leaving 333 SG, 256 G, 414 WNB, and 499 BBH signals to use in our study. Improvements in BW’s detection efficiency were made as a result of this study, details of which are described in Section A in the Appendix. To allow comparison of our results with the ones presented in [Essick et al. \(2015\)](#) and to truly characterize BW’s performance during O1 we used the O1 version of BW, and we excluded these problematic signals to avoid the domination of this feature in our analysis. We present the results of [Essick et al. \(2015\)](#) for the same subset of events we used in this study, to enable a fair comparison of results (see Figure Sets 1-4.).

Results presented here depend on the parameter distributions of injected signals defined in Table 4 of [Essick et al. \(2015\)](#), and on the corresponding detection efficiencies of the combination of cWB and BW pipelines for the different parameter sets. Results are particularly dependent on the network signal-to-noise ratio (SNR_{net}) distribution of injected signals (see inset of Figure 5). However, this SNR_{net} distribution is a good approximation for generic burst signals uniformly distributed in volume (see Appendix C in [Essick et al. 2015](#)).

3.1. Sky localization

BW computes a skymap defined as the posterior probability density function of the GW source location expressed as a function of celestial coordinates α (right ascension) and δ (declination), denoted by $p_{\text{sky}}(\alpha, \delta)$. Example skymaps for each morphology are shown in Appendix B. Skymaps for all the injections can be found in the Burst First2Years sky localization Open Data release¹. There are many possible quantitative measures for the “goodness” of source localization, here we implement the ones defined in [Essick et al. \(2015\)](#).

The first measure is the *angular offset* ($\delta\theta$), which is

the angular distance between the maximum of p_{sky} and the true location of the injected signal. Figure 1 shows normalized histograms of $\cos(\delta\theta)$ for all injections, with the upper axis showing the corresponding $\delta\theta$ values. The distribution has a peak at $\cos(\delta\theta) = 1$, which suggests that BW tends to reconstruct the most probable location of the source close to the actual source location. There is also a smaller peak at $\cos(\delta\theta) = -1$, which indicates that it is more likely that BW reconstructs the opposite direction of the sky compared to the location of the injected signal than a direction perpendicular to the injected signal’s location. This is due to the fact that opposite directions cannot be distinguished using the network antenna pattern which has the same value at opposite directions because of the near co-alignment of H1 and L1 detectors ([Singer et al. 2014](#)). However, the peak at $\cos(\delta\theta) = -1$ is smaller than the one at $\cos(\delta\theta) = 1$ because opposite directions are only allowed by the triangulation ring when the source is right above (or below) the detectors, and thus the triangulation ring is a great circle on the celestial sphere. Note that the distributions for different morphologies are very similar to each other, which means that the angular offset depends weakly on signal morphology. We show summary statistics of $\delta\theta$ distributions for all morphologies in Table 1. It is clearly visible that BW performs best for *BBH* signals, while *SG*, *G* and *WNB* signals show slightly larger $\delta\theta$ values. Statistical errors on reported values are in the order of a few percent. Figure Set 1 shows normalized histograms of $\cos(\delta\theta)$ obtained with the cWB and LIB pipelines on the subset of signals detected by BW.

EM follow-up observations tend to target the point of the sky with the highest p_{sky} value first, and continue with points having lower p_{sky} values. This motivates the introduction of the *searched area* (\mathcal{A}) as a second measure, which is the total sky area observed before aiming a hypothetical telescope at the true location of the source:

$$\mathcal{A} = \int H(p_{\text{sky}}(\alpha, \delta) - p_0) d\Omega \quad (4)$$

where H is the Heaviside step function, p_0 is the value of p_{sky} at the true location of the source, and $d\Omega = \cos \delta d\delta d\alpha$.

We show the cumulative histogram of \mathcal{A} for all injections in Figure 2. Histograms for different morphologies follow a similar trend, but the curves are shifted along the horizontal axis. This can be quantified e.g. with the fraction of recovered signals with $\mathcal{A} < 100 \text{ deg}^2$, which is 29.3% for *G*, 41.3% for *WNB*, 46.5% for *SG*, and 50.1% for *BBH* signals. Another difference between morphologies is that there is a fraction of *WNB* signals with searched area equal to the whole sky ($\mathcal{A} \simeq 4 \cdot 10^4 \text{ deg}^2$). This is due to the fact that $p_0 = 0$ for these signals,

¹ <http://www.ligo.org/scientists/burst-first2years/>

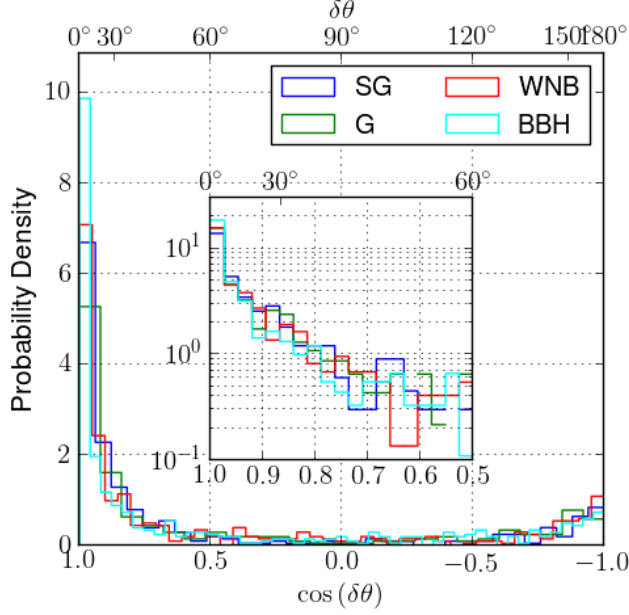


Figure 1. Normalized histograms of *angular offsets* ($\delta\theta$) for injections with four different morphologies (*SG*, *G*, *WNB*, *BBH*). Most of the injected signals have $\cos(\delta\theta) = 1$, which indicates that BW tends to place the most probable location close to the true location. Note that the distributions for different morphologies are very similar to each other, which means that the angular offset does not depend strongly on signal morphology. The complete figure set (3 figures) showing the same plot for cWB and LIB pipelines is available in the online journal.

i.e. the posterior distribution has no support at the true location of the source. There are no such signals with *SG*, *G* and *BBH* morphologies. A reference curve labeled with *SG* (LIB) shows results for the LIB pipeline on the subset of *SG* signals detected by BW. Note that LIB uses a single sine-Gaussian to reconstruct the signal, so for *SG* injections LIB becomes a matched-filtering analysis for which better performance is expected, while BW sometimes uses more than one sine-Gaussian, because it favors more complex signals. It shows that LIB performed similarly, but slightly better for *SG* signals. We show summary statistics of \mathcal{A} distributions for all morphologies in Table 1. It is clearly visible that BW performs best for *BBH* signals, while *SG*, *G* and *WNB* signals show significantly larger \mathcal{A} values. Statistical errors on reported values are in the order of a few percent. Figure Set 2 shows normalized histograms of \mathcal{A} obtained with the cWB and LIB pipelines on the subset of signals detected by BW.

Even if $\delta\theta$ and \mathcal{A} are small, the favored sky positions can still be either well localized or spread out over various parts of the sky. To quantify this feature, we introduce the *extent* ($\delta\theta_{\text{inj}}$) of a skymap as the maximum angular distance between the location of the injected signal

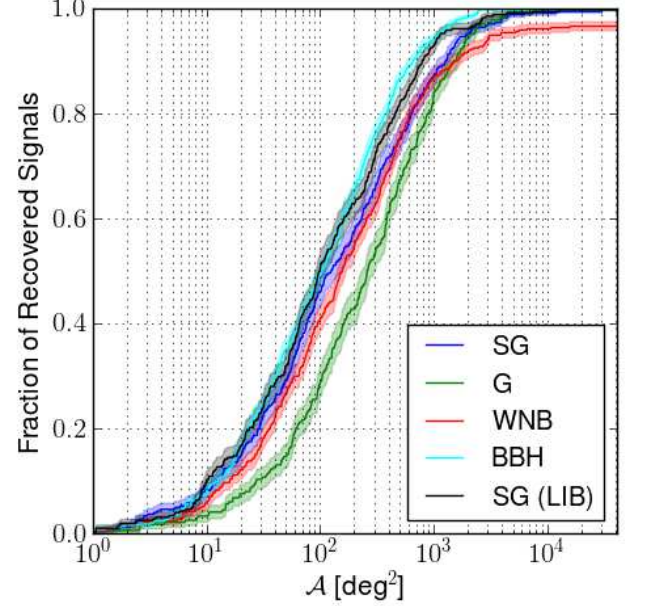


Figure 2. Cumulative histograms of *searched area* (\mathcal{A}). Histograms for different morphologies follow a similar trend, except that the curves are shifted along the horizontal axis. A reference curve labeled with *SG* (LIB) shows results for the LIB pipeline on the subset of *SG* signals detected by BW. The complete figure set (3 figures) showing the same plot for cWB and LIB pipelines is available in the online journal.

and any other point satisfying $p_{\text{sky}}(\alpha, \delta) \geq p_0$. We show histograms of $\delta\theta_{\text{inj}}$ in Figure 3. The shown distributions are clearly bimodal with peaks at $\cos(\delta\theta_{\text{inj}}) = \pm 1$. The peak at $\cos(\delta\theta_{\text{inj}}) = 1$ corresponds to well localized signals, while the peak at $\cos(\delta\theta_{\text{inj}}) = -1$ shows that there is a similarly large fraction of events with the skymap extended even to the opposite direction of the sky compared to the true location of the signal. This is due to the same effect described previously when explaining Figure 1. Note that there are significant differences in the height of the two peaks, e.g. histogram for *BBH* signals have twice as high peak at $\cos(\delta\theta_{\text{inj}}) = 1$ than the histogram for *G* signals. Figure Set 3 shows histograms of $\delta\theta_{\text{inj}}$ obtained with the cWB and LIB pipelines on the subset of signals detected by BW.

Even if previous measures indicate a well localized source, the skymap can still be fragmented, which makes it more difficult to cover the whole with EM observations. We thus introduce the *fragmentation* of a skymap as the number of disjoint regions in the union of points satisfying $p_{\text{sky}}(\delta, \alpha) \geq p_0$. We show the distribution of the number of disjoint regions in Figure 4. Number of disjoint regions is less than 4 for more than 50% of injected signals for all morphology. Skymaps for *SG* and *WNB* signals are significantly more fragmented than for *G* and *BBH* signals. This is due to the fact that the skymaps of these signals are more likely to have “fringe

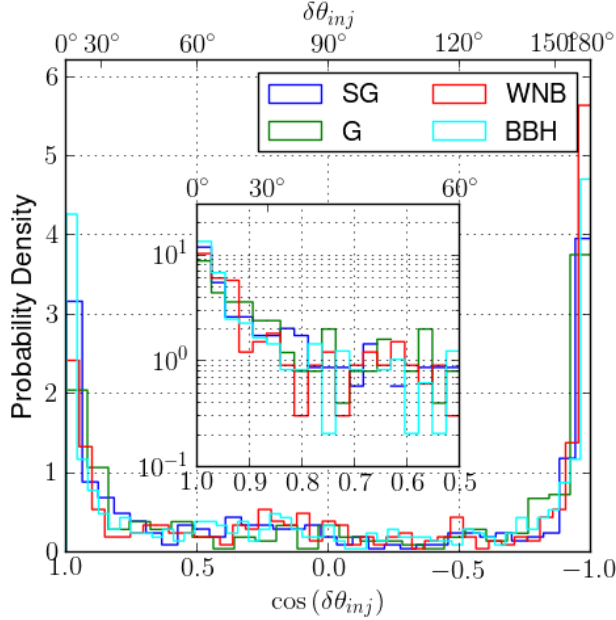


Figure 3. Normalized histograms of the *extent* ($\delta\theta_{inj}$) of skymaps for the four different injection morphologies. The shown distributions are bimodal for all morphologies with peaks at $\cos(\delta\theta_{inj}) = \pm 1$. The complete figure set (3 figures) showing the same plot for cWB and LIB pipelines is available in the online journal.

peaks”. These are separate rings in the sky corresponding to local maxima of matches between different data streams obtained when they are shifted by half-integer multiples of the period of the signal (for details see Section A in the Appendix). Figure Set 4 shows distributions of the number of disjoint regions obtained with the cWB and LIB pipelines on the subset of signals detected by BW.

To compare BW’s performance with LIB’s and cWB’s (Essick et al. 2015), we created the equivalents of Figures 1-4 with LIB and cWB using the same subset of events we used in this study (see Figure Sets 1-4). We have found that all metrics show that these algorithms perform similarly in localizing the source. Histograms of \mathcal{A} show that \mathcal{A} values for BW are comparable to, but systematically bigger than for cWB and LIB for all morphologies, except for *BBH* signals, for which BW typically yields smaller searched areas than LIB. Also, there are more *WNB* skymaps with large searched areas ($\mathcal{A} \gtrsim 100 \text{ deg}^2$) for LIB than for BW. This is likely due to its ability to recover more of the signal by using multiple wavelets as opposed to a single sine-Gaussian template.

3.2. Waveform reconstruction

BW uses sine-Gaussian wavelets to reconstruct a GW signal from the detector output, which means that the recovered signal is always given as a linear combination

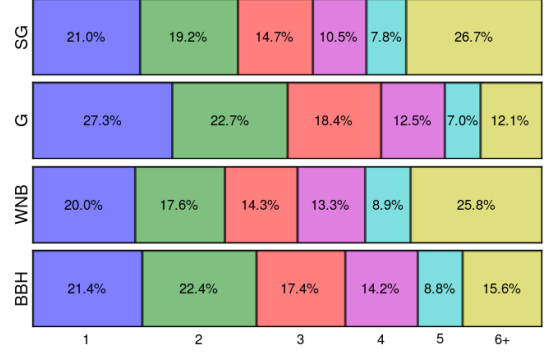


Figure 4. Distributions of *fragmentation*. Each row corresponds to one of the four morphologies (*SG*, *G*, *WNB*, *BBH*). Numbers at the bottom of the chart represent the number of disjoint regions in parts of the sky where $p_{sky} \geq p_0$. Number of disjoint regions is less than 4 for more than 50% of injected signals for all morphologies. The complete figure set (3 figures) showing the same plot for cWB and LIB pipelines is available in the online journal.

of sine-Gaussian wavelets, the number of which is a parameter in the RJMCMC. To characterize the quality of waveform reconstruction, we introduce the *overlap* (\mathcal{O} , sometimes referred to as *match*) which measures the similarity of an injected (h_i) and a recovered (h) waveform as:

$$\mathcal{O} = \frac{(h_i|h)}{\sqrt{(h_i|h_i)(h|h)}}, \quad (5)$$

where $(\cdot|\cdot)$ is a noise weighted inner product, defined as:

$$(a|b) = 2 \int_0^\infty \frac{a(f)b^*(f) + a^*(f)b(f)}{S_n(f)} df, \quad (6)$$

where S_n is the one-sided power spectral density of the detector noise, and x^* denotes the complex conjugate of x .

From Eq. (5) it is visible that \mathcal{O} ranges from -1 to 1, with $\mathcal{O} = 1$ meaning perfect match between h_i and h , $\mathcal{O} = 0$ meaning no match at all, and $\mathcal{O} = -1$ meaning a perfect anti-correlation between h_i and h . With Eq. (5), we can calculate the overlap using data from only one detector. To characterize the waveform reconstruction for the network of GW detectors, we introduce the *network overlap* (\mathcal{O}_{net}) by changing the inner products in Eq. (5) with the sum of inner products calculated for different detectors:

$$\mathcal{O}_{net} = \frac{\sum_{j=1}^N (h_i^{(j)}|h^{(j)})}{\sqrt{\sum_{j=1}^N (h_i^{(j)}|h_i^{(j)}) \cdot \sum_{j=1}^N (h^{(j)}|h^{(j)})}}, \quad (7)$$

where j denotes the j -th detector in the network, and N is the number of detectors used in the analysis (note that $N = 2$ in this study).

Table 1. Summary statistics of \mathcal{A} and $\delta\theta$ distributions. Statistical errors are in the order of a few percent.

morphology		BBH	SG	G	WNB
fraction (in %) with searched area less than	5 deg ²	3.6	4.8	2.3	2.7
	20 deg ²	17.4	15.6	7.8	12.3
	100 deg ²	50.1	46.5	29.3	41.3
	200 deg ²	66.5	58.6	43.4	56.0
	500 deg ²	87.0	75.4	67.6	76.1
	1000 deg ²	94.6	87.7	84.4	87.4
fraction (in %) with $\delta\theta$ less than	1°	3.0	1.2	1.6	1.0
	5°	15.4	10.5	12.1	10.1
	15°	37.5	31.2	30.9	30.2
	45°	62.7	69.1	62.9	61.4
	60°	69.1	75.7	68.4	67.1
	90°	76.4	79.9	75.4	76.1
median searched area		99.2 deg ²	121.3 deg ²	252.8 deg ²	151.0 deg ²
median $\delta\theta$		25.1°	26.2°	29.9°	30.3°

Figure 5 shows the fraction of recovered signals with \mathcal{O}_{net} greater than a given value. Shaded ranges represent the 1σ uncertainty calculated using the Dvoretzky–Kiefer–Wolfowitz inequality (Dvoretzky et al. 1956). The fraction of injected signals with $\mathcal{O}_{\text{net}} > 0.9$ is 98% for G , 96% for SG , 47% for BBH , and 45% for WNB signals after the waveform reconstruction with BW. 95% of injections have $\mathcal{O}_{\text{net}} > 0.92$ for G signals, $\mathcal{O}_{\text{net}} > 0.91$ for SG signals, $\mathcal{O}_{\text{net}} > 0.75$ for BBH signals, and $\mathcal{O}_{\text{net}} > 0.68$ for WNB signals. In Figure 5, the higher the curves reach at a given \mathcal{O}_{net} value, the better the reconstruction is. This suggests that BW’s waveform reconstruction works most effectively for SG and G signals, for which the curves are identical within the 1σ statistical error. BW’s waveform reconstruction is less effective for WNB and BBH signals, and it shows similar characteristics for these morphologies at high network overlaps ($\gtrsim 0.8$), but the distribution for WNB signals has a longer tail at low \mathcal{O}_{net} values. The better performance of BW for SG and G signals is due to the fact that at low SNR_{net} BW tends to use fewer wavelets to avoid overfitting the data. SG and G signals can be reconstructed accurately even with just 2-3 sine-Gaussian wavelets, while this is not possible for WNB and BBH signals. This also means that the curves for SG and G signals in Figure 5 represent BW’s maximal capability of reconstructing a GW signal for a given noise level, while the results for WNB and BBH signals represent BW’s performance on more generic (and thus, more realistic) GW signals. Note that while \mathcal{O}_{net} values are smaller for WNB and BBH signals, BW detects these with more confidence, because its detection statistic has a stronger dependence on signal complexity than on SNR_{net} (for

details see Littenberg et al. 2016). The inset plot in Figure 5 shows the normalized histogram of injected signals’ network signal-to-noise ratio (SNR_{net}) for the four different signal morphologies. SG and G signals have an overabundance at $\text{SNR}_{\text{net}} \lesssim 20$ relative to WNB and BBH signals. This indicates that the previously described difference in the distribution of \mathcal{O}_{net} is not due to the different SNR_{net} distributions, as BW performs better for SG and G signals despite the fact that SNR_{net} values for SG and G signals are usually smaller than for WNB and BBH signals. Note that these distributions strongly depend on the parameter distributions of injected signals defined in Table 4 of Essick et al. (2015), and on the corresponding detection efficiencies of the combination of cWB and BW pipelines for the different parameter sets (see the SNR_{net} histogram in the inset of Figure 5).

We show \mathcal{O}_{net} vs. SNR_{net} for SG , G , and WNB signals in the left panel of Figure 6, where shaded regions between dashed lines represent the 1σ uncertainty regions calculated with the bootstrap method. Curves were estimated from \mathcal{O}_{net} and SNR_{net} values with a Gaussian kernel smoother, which is a nonparametric regression method. Note that we excluded the injections with $\text{SNR}_{\text{net}} > 100$ from the estimation of these curves, and we only show the estimated curves up to $\text{SNR}_{\text{net}}=70$. All three morphologies show a clear trend of \mathcal{O}_{net} increasing with SNR_{net} .

For BBH signals we calculated the \mathcal{O}_{net} vs. SNR_{net} curves in two separate bins of total mass (M_{tot}) of the binary black hole system, calculated in the detector’s frame. The two bins were defined with M_{tot} being $M_{\text{tot}} < \hat{M}_{\text{tot}}$ and $M_{\text{tot}} > \hat{M}_{\text{tot}}$, where $\hat{M}_{\text{tot}} = 44.49 M_{\odot}$

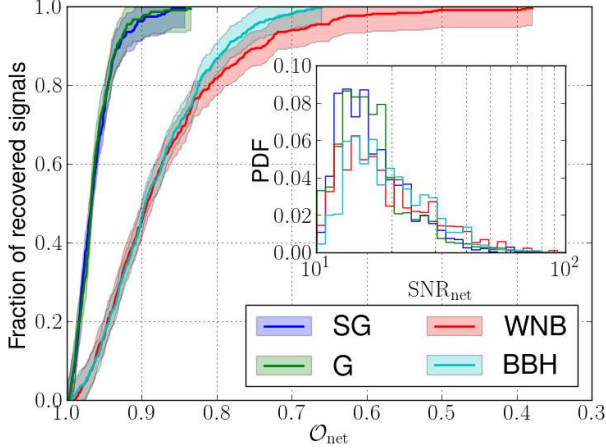


Figure 5. Fraction of recovered signals, with overlaps exceeding a value given in the horizontal axis (note that the axis is reversed). Shading represents the 1σ uncertainty calculated using the Dvoretzky–Kiefer–Wolfowitz inequality. The higher the curves reach at a given network overlap value, the better the reconstruction is. The inset shows the normalized histogram of injected signals’ network signal-to-noise ratio (SNR_{net}) for the four different morphologies. The curves for *SG* and *G* signals are the same within the 1σ statistical error, and they indicate significantly better reconstruction of *SG* and *G* signals than *WNB* and *BBH* signals.

is the median of M_{tot} values for all *BBH* injections. The \mathcal{O}_{net} vs. SNR_{net} curves for *BBH* signals are shown in the right panel of Figure 6. Similarly to other morphologies, *BBH* injections also show a clear trend of increasing \mathcal{O}_{net} with increasing SNR_{net} . At low ($\lesssim 35$) SNR_{net} values, BW performed significantly better for signals with higher M_{tot} , while differences in the curves are within the level of statistical errors for higher SNR_{net} values. Signals with high M_{tot} are recovered with better accuracy because a large fraction of the signal power is in a compact region of time-frequency space and therefore can be captured with a small number of wavelets, while signals with low M_{tot} spend a comparatively longer amount of time in the sensitive band of the detectors, requiring more wavelets, and a larger total signal strength, to achieve a similar fit. This difference vanishes at high SNR_{net} because BW uses more wavelets to reconstruct signals with higher SNR_{net} .

Similarly to Figure 5, Figure 6 also shows that BW performs very similarly on *SG* and *G* signals, and much less efficiently on *WNB* and *BBH* signals. This is due to the fact that BW needs to use more wavelets to accurately reconstruct *WNB* and *BBH* signals. Note, that despite the weaker performance on *WNB* and *BBH* signals, these also approach the reconstruction accuracy for *SG* and *G* signals at higher SNR_{net} values. Comparing the two panels of Figure 6, it is visible that the curve for *BBH* signals is similar to the curve for *WNB* signals, with slightly worse overlap at low SNR_{net} and slightly

better overlap at high SNR_{net} values.

Our results show that BW robustly reconstructs waveforms with various morphologies. Although there are significant differences between the efficiency of reconstructions of signals with different morphologies, even for the worst case of *WNB* signals (which do not even match BW’s assumption of the signal always being elliptically polarized), most of them have relatively high overlaps, and there is a clear trend of \mathcal{O}_{net} approaching 1 as SNR_{net} increases.

3.3. Point estimates of waveform moments

For a generic burst signal, we do not have any specific astrophysical model whose parameters could be estimated. In this case we can still give estimates on model-independent parameters of the signal. Here we consider the moments of the waveform as such parameters.

The first moments are central time (t_0) and central frequency (f_0), and the second moments are duration (Δt) and bandwidth (Δf), defined as

$$t_0 = \int_{-\infty}^{\infty} dt \rho_{\text{TD}}(t)t \quad (8a)$$

$$f_0 = \int_0^{\infty} df \rho_{\text{FD}}(f)f \quad (8b)$$

$$(\Delta t)^2 = \int_{-\infty}^{\infty} dt \rho_{\text{TD}}(t)(t - t_0)^2 \quad (8c)$$

$$(\Delta f)^2 = \int_0^{\infty} df \rho_{\text{FD}}(f)(f - f_0)^2 \quad (8d)$$

respectively, where ρ_{TD} and ρ_{FD} are the effective normalized distributions of signal energy, expressed in the time domain (TD) and in the frequency domain (FD):

$$\rho_{\text{TD}} = \frac{h_{\times}^2(t) + h_{+}^2(t)}{h_{\text{rss}}^2}, \quad (9a)$$

$$\rho_{\text{FD}} = \frac{2(|\tilde{h}_{\times}(f)|^2 + |\tilde{h}_{+}(f)|^2)}{h_{\text{rss}}^2}. \quad (9b)$$

Estimations of higher order moments could also be given with BW, however we excluded them from our analysis due to the fact that they are more strongly affected by statistical errors than estimations of the first order moments (for a detailed discussion of this, see the end of this section).

BW reconstructs the waveform and calculates the waveform moments for each sample in the Markov chain. We calculated the median value to give a point estimate of the waveform moments. To quantify the accuracy of

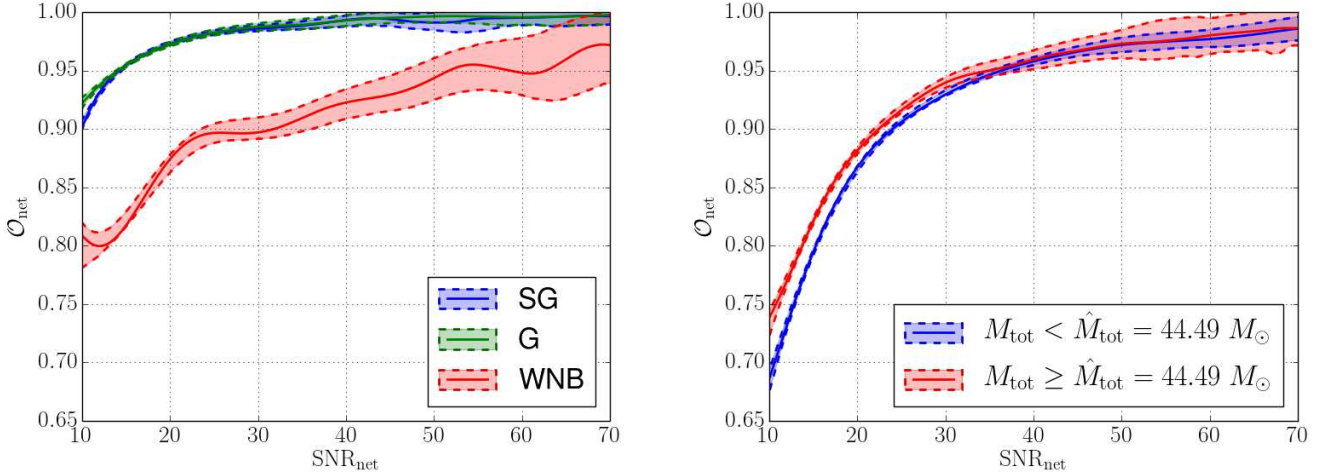


Figure 6. Dependence of network overlaps (\mathcal{O}_{net}) on network signal-to-noise ratios (SNR_{net}) for *SG*, *G*, *WNB*, and *BBH* signals. Note that we excluded the injections with $\text{SNR}_{\text{net}} > 100$ from the curve estimation. Shaded areas represent the 1σ uncertainty regions of the measured \mathcal{O}_{net} values. Left panel shows SNR_{net} dependence of \mathcal{O}_{net} for *SG*, *G*, and *WNB* signals. All three morphologies show a clear trend of increasing overlap with increasing SNR_{net} . Right panel shows SNR_{net} dependence of network overlaps for *BBH* signals with detector-frame total mass below and above the median total mass $\hat{M}_{\text{tot}} = 44.49 M_{\odot}$. BW performed significantly better for signals with higher M_{tot} at $\text{SNR}_{\text{net}} \lesssim 35$ values.

the point estimate of waveform moment x , we define the absolute error of the estimation, e_x , as:

$$e_x = |x^{(e)} - x^{(r)}|, \quad (10)$$

where $x^{(e)}$ is the estimated, and $x^{(r)}$ is the real value of x . We also introduce the relative error of an estimate, η_x , as:

$$\eta_x = \frac{e_x}{x^{(r)}}. \quad (11)$$

We show the probability density functions of e_{t_0} , $\eta_{\Delta t}$, η_{f_0} , and $\eta_{\Delta f}$ in Figure 7, all estimated with kernel density estimation. All moments were calculated for H1 detector data, however, results are very similar for the L1 detector too. We chose to show e_{t_0} instead of η_{t_0} because only absolute time differences have physical meaning in GW detections. Vertical dashed lines in Figure 7. indicate the 95th percentile of distributions, i.e. 95% of injected signals had smaller errors than these values.

Probability density functions of e_{t_0} for different morphologies are shown in the top left panel of Figure 7. This shows that the most accurate t_0 estimates with BW are obtained for *SG* signals, while estimates for *BBH* signals are the least accurate. e_{t_0} values are in the same order of magnitude as the 10 ms light travel time between H1 and L1, but this doesn't affect sky localization because e_{t_0} values we obtain for H1 and L1 are strongly correlated. The relatively large mismatch between the estimated and real values of t_0 is due to the fact we can miss parts of the signal in both detectors.

For example for *BBH* signals BW is almost insensitive to the inspiral of the waveform, which makes up the bulk of the signal duration. This bias gets progressively worse for lower mass systems. This also explains why the estimation of t_0 is less accurate for *BBH* signals, because this effect is smaller for the other three morphology. Note that even *BBH* signals have $e_{t_0} < 17.5$ ms for 95% of the injected signals.

Probability density functions of $\eta_{\Delta t}$ are shown in the top right panel of Figure 7. The $\eta_{\Delta t}$ curves are significantly different for different waveform morphologies, and they are less affected (compared to e_{t_0}) by the different parameter distributions of injected signals. For most of the *SG* and *G* signals, and for a significant fraction of *WNB* signals, relative errors less than 20% are achievable, while this only holds for a moderate fraction of *BBH* signals. Estimation on *G* signals have less than 25% error for 95% of injected signals, and even for *BBH* signals the 95th percentile is 56%.

Probability density functions of η_{f_0} are shown in the bottom left panel of Figure 7. Similarly for the Δt estimations, f_0 estimations are the most accurate for *SG* signals and the least accurate for *BBH* signals. For 95% of *SG* signals, f_0 can be estimated with a less than 3.4% error, and even for *BBH* signals η_{f_0} is less than 17% for 95% of injected signals.

Probability density functions of $\eta_{\Delta f}$ are shown in the bottom right panel of Figure 7. The accuracy of Δf estimation seems to be similar for different morphologies, with 95th percentiles varying between 20% and 40%. It

is important to notice that relative errors of bandwidth estimations tend to be higher than of central frequency estimations. This is due to the fact that second order moments inherit errors from estimations of lower order moments (see Eq. (8c) and (8d)), and thus have higher statistical errors. We expect that third and higher order moments would have even bigger errors, and that is why we only examined the first two moments. Medians and 95th percentiles of errors for each moment and morphology are shown in Table 2.

These results show that the distributions of errors are very similar for different morphologies, the only exceptions are the t_0 and Δt estimates for *BBH* signals which show a much flatter distribution than any other errors examined for any morphology, and also lack clear peaks around zero error. There is a tendency among errors of the four estimated moments that estimations for *SG* and *G* signals are more precise than estimations for *WNB* and *BBH* signals. This is the consequence of the fact that *SG* and *G* signals have higher overlaps than *WNB* and *BBH* signals (see Figure 5). There are, however, deviations from this tendency, e.g. the median and 95th percentile of η_{f_0} is higher for *G* signals than for *WNB* signals (see Table 2), despite the fact that *G* signals have significantly higher overlaps (see Figure 5). This means that it is not overlap alone that influences errors of estimated waveform moments, but they are influenced by the morphology of the signal as well.

4. CONCLUSION

We presented a comprehensive multi-aspect study on the performance of BW, a Bayesian GW burst PE pipeline used by the LIGO-Virgo Collaboration for reconstructing GW burst signals and their parameters. We injected a large number of simulated signals with four different morphologies (sine-Gaussians, Gaussians, white-noise bursts, and binary black hole signals) into simulated O1 aLIGO noise to test BW's performance in three different aspects of PE: sky localization, waveform reconstruction, and estimation of waveform moments (for details on the methods we used see Section 2).

BW localizes sources with a level of accuracy comparable for all four morphologies, with the median separation of actual and estimated sky locations ranging from 25.1° to 30.3° (see Table 1), and median searched area (\mathcal{A} , see Eq. 4) ranging from 99.2 deg^2 to 252.8 deg^2 (see Section 3.1). This is reasonable accuracy for a two-detector network, and is comparable to accuracies of other localization pipelines (cWB and LIB) studied previously (Essick et al. 2015). Histograms of \mathcal{A} (see Figure 2) show that \mathcal{A} values for BW are comparable to, but systematically bigger than for cWB and LIB for all morphologies. The exceptions are *BBH* signals, for

which BW's \mathcal{A} values are systematically smaller. Note that the runtime of cWB and LIB is much shorter than of BW.

BW reconstructs waveforms as a linear combination of sine-Gaussian wavelets. To measure the goodness of reconstruction, we used the network overlap (\mathcal{O}_{net} , see Eq. 7), which quantifies the similarity between the injected and the reconstructed signals. We have found that BW reconstructs signals with $\mathcal{O}_{\text{net}} > 0.9$ for 98% of *G*, 96% of *SG*, 45% of *WNB*, and 47% of *BBH* signals (see Section 3.2). We have also found that (see Figure 6) \mathcal{O}_{net} increases rapidly with increasing SNR_{net} , reaching $\mathcal{O}_{\text{net}} = 0.95$ at $\text{SNR}_{\text{net}} \approx 14$ for *SG* and *G*, at $\text{SNR}_{\text{net}} \approx 50$ for *WNB*, and at $\text{SNR}_{\text{net}} \approx 35$ for *BBH* signals. These results suggest that we can expect very good reconstruction ($\mathcal{O}_{\text{net}} > 0.95$) for almost any signal with high ($\gtrsim 50$) SNR_{net} , and reasonably good reconstruction ($\mathcal{O}_{\text{net}} > 0.85$) for almost any signal with moderate ($\gtrsim 20$) SNR_{net} .

We also examined how precisely BW can estimate the waveform moments of a GW signal (see Section 3.3). These are model-independent parameters of a signal, therefore by examining the estimation of them, we can characterize PE without assuming any astrophysical model for the source. We have found that the accuracy of these estimations depend on the \mathcal{O}_{net} distribution of signals, but the morphology of the signal also affects them (see Figure 7). The median relative error on the estimation of central frequency is 0.67% for *SG*, 3.01% for *G*, 2.09% for *WNB*, and 4.28% for *BBH* signals (see Table 2). There is no standard procedure on how the estimated moments of GW bursts can be used to test astrophysical models, however future studies can use our results to test the feasibility of particular methods using signal moments.

This paper fits into a series of studies examining PE for GW bursts (see e.g. Klimenko et al. 2011; Essick et al. 2015). These studies can be used in comparisons with improved performances of future PE pipelines, and in testing the feasibility of possible astrophysical applications of future GW burst detections.

This paper was reviewed by the LIGO Scientific Collaboration under LIGO Document P1600181. We thank Marco Drago and Sergey Klimenko for their valuable comments on the manuscript. We acknowledge the Burst First2Years sky localization Open Data release². The authors acknowledge the support of the National Science Foundation and the LIGO Laboratory. LIGO was constructed by the California Institute of Tech-

² <http://www.ligo.org/scientists/burst-first2years/>

Table 2. Medians (50th percentiles) and 95th percentiles of waveform moment errors for the *SG*, *G*, *WNB*, and *BBH* signal morphologies. *P* denotes the percentile rank of values given in the corresponding table columns.

	<i>P</i>	Signal morphology			
		<i>SG</i>	<i>G</i>	<i>WNB</i>	<i>BBH</i>
e_{t_0} [ms]	50th	0.47	0.32	0.93	6.90
	95th	1.79	1.20	6.59	17.45
$\eta_{\Delta t}$ [%]	50th	1.17	2.43	5.04	21.45
	95th	25.28	39.75	42.21	56.98
η_{f_0} [%]	50th	0.67	3.01	2.09	4.28
	95th	3.34	11.03	8.41	16.91
$\eta_{\Delta f}$ [%]	50th	5.68	6.85	6.84	5.55
	95th	22.82	20.90	39.19	29.60

nology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058. The authors would like to acknowledge the use of the LIGO Data Grid computer clusters for performing all the computation reported in the paper. Bence Bécsey was supported by the ÚNKP-16-2 New National Excellence Program of the Ministry of Human Capacities. Bence Bécsey

was supported by the Hungarian Templeton Program that was made possible through the support of a grant from Templeton World Charity Foundation, Inc. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of Templeton World Charity Foundation, Inc. Peter Raffai is grateful for the support of the Hungarian Academy of Sciences through the "Bolyai János" Research Scholarship programme.

APPENDIX

A. RESOLVING BAYESWAVE O1 VERSION'S ISSUE WITH HIGH-*Q* SIGNALS

During O1, BW was prone to classifying simulated short-duration high-frequency signals which underwent many wave cycles (i.e. high-*Q* signals) while in the measurement band of the detector as glitches. In principle there is no reason for the Bayesian evidence used to rank hypothesis under consideration by BW to have strong frequency dependence.

Upon examination of the mis-classified injections, it was revealed that the high-*f*, high-*Q* signals exhibit multimodal likelihood support in the $(\alpha, \delta, t_0, f_0)$ parameter sub-space. For these signals, the Markov Chain Monte Carlo (MCMC) sampler, which serves as the central engine to the BW algorithm, was not generically sampling between the different modes, and was thus prone to missing significant portions of the coherent signal and preferring the incoherent glitch model (which does not suffer the correlations between time-frequency parameters and sky location).

The cause of the multimodal likelihood function is clear. For a sinusoidal signal ($Q = \infty$) the waveform is perfectly degenerate when time-shifted by an integer number of wave-periods (T). For high-*Q* signals, a number of integer periods (or half-integer periods with a π radians phase shift), time shifts produce similarly good fits to the data. For coherent signals, these (nearly) degenerate time shifts are also present in the time delay between detectors, which, for BW, is encoded in the sky location.

To overcome BW's susceptibility to missing modes of the likelihood when analyzing high-*Q* signals, we added a proposal distribution to the MCMC which explicitly suggests half-integer-period time shifts, along with half-integer-cycle phase shifts, for the wavelet parameters. Furthermore, extensive development (beyond the scope of this paper) to improve the overall capabilities of BW's MCMC to sample the complicated sky-location posteriors encountered by two-detector gravitational-wave networks has been completed.

Figure 8 contains two scatter plots from the BW MCMC utilizing the dedicated proposal distributions. The multimodal nature of the posterior is on clear display, as is the efficiency with which the MCMC sampler is able to move between local maxima in the likelihood. This example came from an $f \sim 512$ Hz, $Q \sim 40$ sine-Gaussian injection.

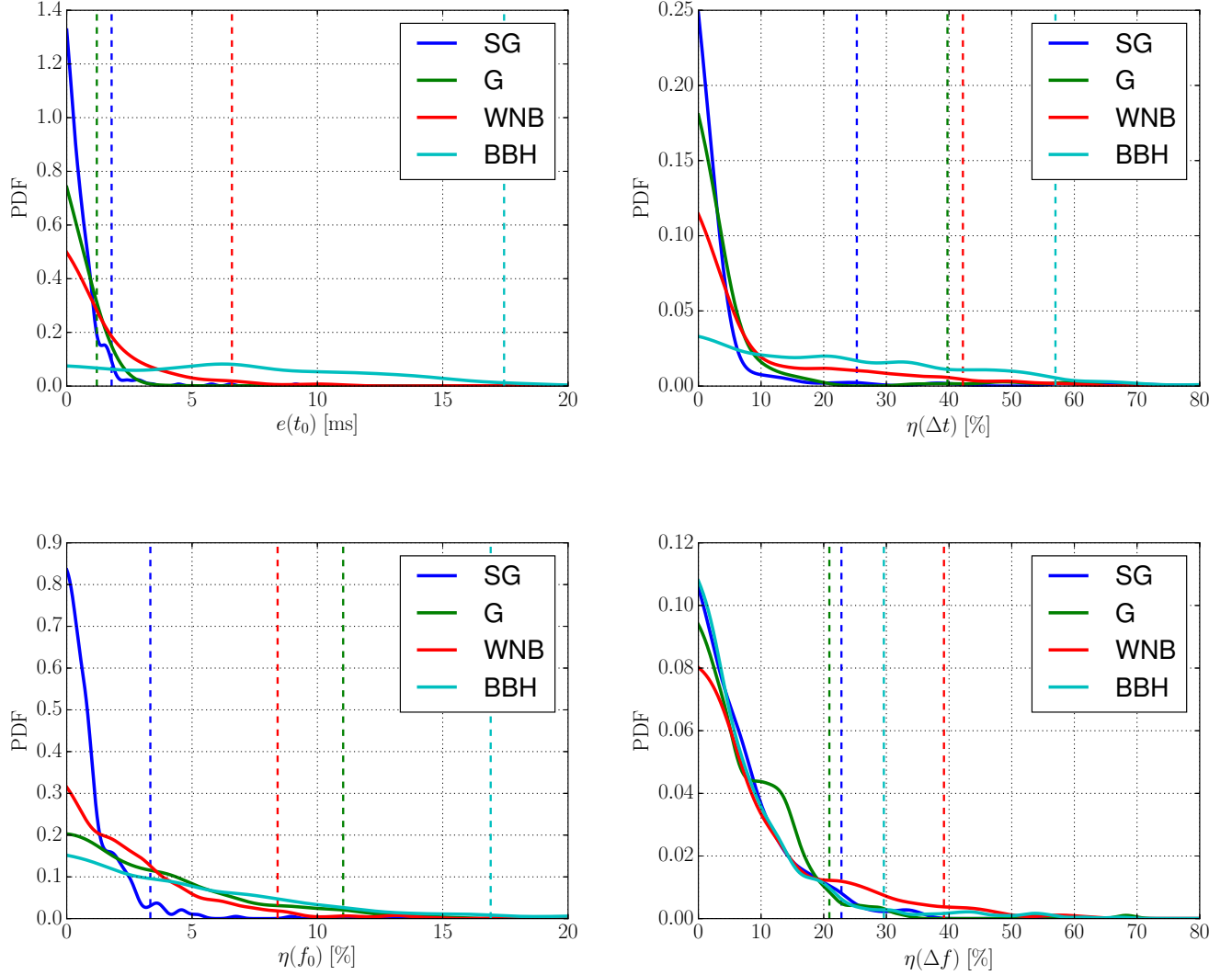


Figure 7. Probability density function (PDF) of waveform moment errors estimated with kernel density estimation. Panels from left to right and from upper to lower show PDFs of absolute errors of central time estimations (e_{t_0}), relative errors of duration estimations ($\eta_{\Delta t}$), relative errors of central frequency estimations (η_{f_0}), and relative errors of bandwidth estimations ($\eta_{\Delta f}$). Colors indicate PDFs for signals with sine-Gaussian (SG), Gaussian (G), white noise burst (WNB), and binary black hole (BBH) morphologies. Vertical dashed lines mark the 95th percentiles of the distributions, marked with the same colors we used for plotting the corresponding PDFs. Values of 95th percentiles and medians can be found in Table 2.

Using the MCMC as it was during O1 we found a preference for the incoherent “glitch” model, with a Bayes factor between that and the coherent “signal” model of $\sim e^{60}$ in favor of the glitch model. Using the updated sampler and analyzing the same (simulated) data we find a Bayes factor of $\sim e^{18}$ in favor of the signal model.

Despite this upgrade to BW’s MCMC engine, we elected to present results *as the algorithm performed* during O1 to facilitate a direct comparison with the snapshot of other burst parameter estimation techniques during the first observing run. Future studies showing how the upgraded sampler performs on similar injections are underway.

B. EXAMPLE SKYMAPS

Figure 9 shows an example skymap for an injected SG signal. The injected location is marked with a star and the corresponding triangulation ring for L1 and H1 detectors is denoted with a grey line. H-L and L-H marks the direction between the two detectors, H+ and L+ the directions above the detectors, and H- and L- the directions below the detectors. Skymap in Figure 9 is a typical one. It is consistent with the triangulation ring of the two detector network and the constraint of the network antenna pattern, which leads to a relatively small elongated area on the sky with

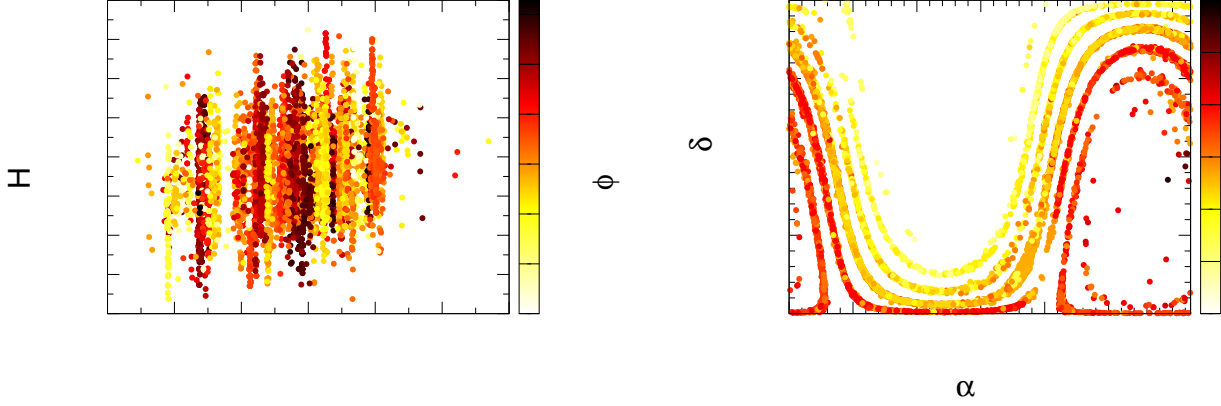


Figure 8. Scatter plot of MCMC samples for signal model parameters of a high- Q , high- f sine-Gaussian injection. The left panel shows the time-frequency plane with points colored by the wavelet phase parameter. Multiple modes and their phase-dependence are evident. The right panel shows the same chain samples, but now projected on the the sky-location plane of the parameter space and colored by the time parameter. Here again it is plain to see how different half-integer-period time shifts correspond to different “rings” on the sky, making this a challenging distribution to sample without well-tuned proposal distributions.

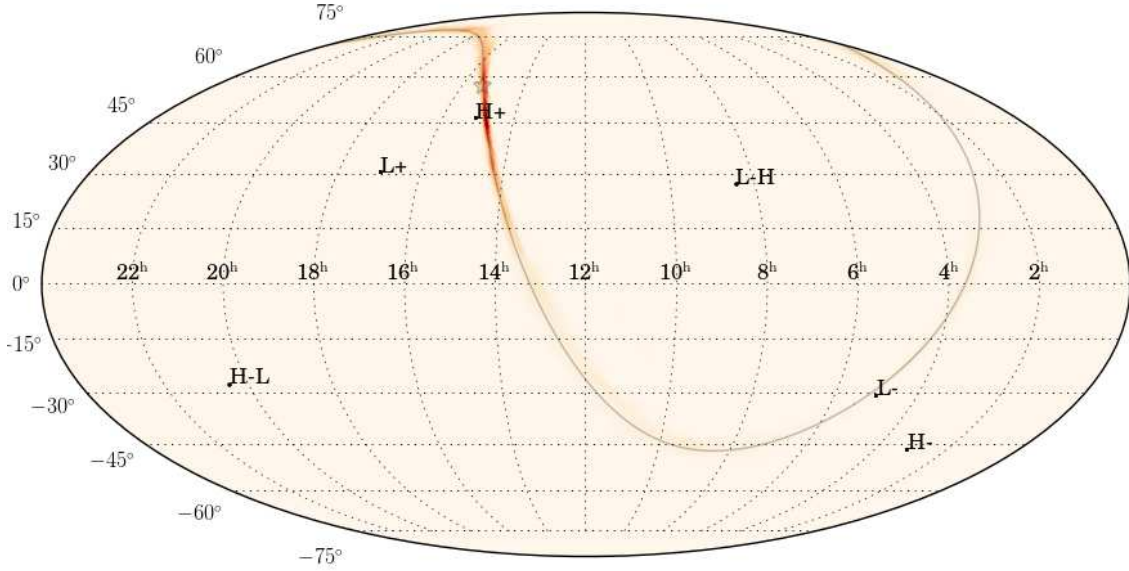


Figure 9. An example skymap showing the reconstructed sky location for an injected SG signal. The injected location is marked with a star and the corresponding triangulation ring for L1 and H1 detectors is denoted with a grey line. H-L and L-H marks the direction between the two detectors, H+ and L+ the directions above the detectors, and H- and L- the directions below the detectors. The complete figure set (20 figures) showing 5 example skymaps for each morphology is available in the online journal. Skymaps for all the signals used in this study are available in the Burst First2Years sky localization Open Data release (<http://www.ligo.org/scientists/burst-first2years/>).

the maximum close to the injected location. Figure Set 8 shows 20 example skymaps (5 for each morphology) in the online journal. Skymaps for all the signals used in this study are available in the Burst First2Years sky localization Open Data release³.

³ <http://www.ligo.org/scientists/burst-first2years/>

REFERENCES

- Aasi, J., Abbott, B. P., Abbott, R., et al. 2015, *Classical and Quantum Gravity*, 32, 074001
- Abadie, J., Abbott, B. P., Abbott, R., et al. 2012, *PhRvD*, 85, 122007
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016a, *ApJL*, 818, L22
- . 2016b, *Physical Review Letters*, 116, 241103
- . 2016c, *ApJL*, 826, L13
- . 2016d, *Physical Review Letters*, 116, 061102
- . 2016e, *PhRvD*, 93, 122004
- . 2016f, *Physical Review Letters*, 116, 241102
- . 2016g, *Living Reviews in Relativity*, 19, arXiv:1304.0670
- Ajith, P., Hannam, M., Husa, S., et al. 2011, *Physical Review Letters*, 106, 241101
- Berry, C. P. L., Mandel, I., Middleton, H., et al. 2015, *ApJ*, 804, 114
- Chatterji, S. 2005, PhD thesis, Massachusetts Institute of Technology
- Cornish, N. J., & Littenberg, T. B. 2015, *Classical and Quantum Gravity*, 32, 135012
- Dvoretzky, A., Kiefer, J., & Wolfowitz, J. 1956, *Ann. Math. Statist.*, 27, 642
- Essick, R., Vitale, S., Katsavounidis, E., Vedovato, G., & Klimenko, S. 2015, *ApJ*, 800, 81
- Green, P. J. 1995, *Biometrika*, 82, 711
- Hannam, M., Husa, S., Ohme, F., Müller, D., & Brüggmann, B. 2010, *PhRvD*, 82, 124008
- Kanner, J. B., Littenberg, T. B., Cornish, N., et al. 2016, *PhRvD*, 93, 022002
- Klimenko, S., Yakushin, I., Mercer, A., & Mitselmakher, G. 2008, *Classical and Quantum Gravity*, 25, 114029
- Klimenko, S., Vedovato, G., Drago, M., et al. 2011, *PhRvD*, 83, 102001
- . 2016, *PhRvD*, 93, 042004
- Littenberg, T. B., Kanner, J. B., Cornish, N. J., & Millhouse, M. 2016, *PhRvD*, 94, 044050
- Lynch, R., Vitale, S., Essick, R., Katsavounidis, E., & Robinet, F. 2015, *ArXiv e-prints*, arXiv:1511.05955
- Singer, L. P., Price, L. R., Farr, B., et al. 2014, *The Astrophysical Journal*, 795, 105
- Veitch, J., Raymond, V., Farr, B., et al. 2015, *Phys. Rev. D*, 91, 042003
- Vitale, S., Essick, R., Katsavounidis, E., Klimenko, S., & Vedovato, G. 2016, *Monthly Notices of the Royal Astronomical Society: Letters*, <http://mnrasl.oxfordjournals.org/content/early/2016/11/30/mnrasl.slw239.f>